引用格式：Ye Pei，Xu Ke，Xu Xiyu. Study on Denoising the Instrumental Errors of the Sea Surface Height Series Derived from Buoys[J]. Remote Sensing Technology and Application，2015，30(4)：661-666.［叶沛，许可，徐曦煜. 基于奇异谱分析的DGPS浮标海面高测量误差研究[J]. 遥感技术与应用，2015，30(4)：661-666.］
doi:10.11873/j.issn.1004-0323.2015.4.0661

基于奇异谱分析的DGPS浮标海面高测量误差研究

叶　沛[1,2,3]，许　可[1,2]，徐曦煜[1,2]
（1. 中国科学院空间科学与应用研究中心微波遥感技术重点实验室，北京　100190；
2. 中国科学院空间科学与应用研究中心，北京　100190；
3. 中国科学院大学，北京　100049）

摘要：在高度计绝对定标实验中，利用差分GPS(Differential GPS，DGPS)浮标进行现场海面高测量得到的海面高(Sea Surface Height，SSH)序列之中，除了包含有由风浪引起的高频噪声外，还包括了仪器误差造成的影响。前者可通过低通滤波方式消除，而后者由于是全频段的白噪声，不能通过低通滤波的方式消除。通过研究奇异谱分析(Singular Spectrum Analysis，SSA)的方法，确定了嵌入维数和截断长度的选取准则，并对其进行仿真，验证了该方法在消除海面高序列的仪器误差噪声中的有效性，同时也验证了嵌入维数和截断长度选取准则的有效性。最后，利用该方法处理了某次DGPS浮标实验获得的海面高序列，处理结果与原始序列相比较为平滑，且在精度上有一定的改善，验证了奇异谱分析方法的有效性。

关　键　词：高度计定标；海面高序列；仪器误差；奇异谱分析；降噪
中图分类号：P 714　　**文献标志码**：A　　**文章编号**：1004-0323(2015)04-0661-06

1　引　言

在高度计定标中，利用差分GPS浮标获得的海面高序列之中通常包含了由风浪、仪器误差等引起的噪声分量，需要对这个原始时间序列进行平滑处理，从而消除噪声分量，得到星下点的平均海面高[1]。为了消除海面高序列中的噪声分量，通常的做法是对该序列进行滑动平均或Vondrak滤波处理以消除其高频分量。然而，分析原始序列频谱可知，由风浪引起的噪声其频谱仅集中在某一频段内，而仪器误差引起的噪声分量则分布于整个频带上。滑动平均或者Vondrak滤波本质上是一种低通滤波，因此可用于消除风浪噪声，但仪器噪声则不能通过这两种方法完全消除。

消除仪器误差常用的方法有最小均方滤波和最小二乘滤波[2]，这两种方法都要求参考输入序列与原始序列相关。最小均方滤波要求输入序列必须是平稳的，而最小二乘滤波是以误差的平方和最小作为最佳准则的滤波方法，可适用于非平稳序列。用该方法进行滤波时，需要知道序列中噪声的统计特性，从而将一个与原始序列中噪声的统计特性相同的序列作为参考输入，实现滤波。

由于一个含有噪声的矩阵的特征值反映了有用信号和噪声的分布情况，因此可用奇异谱分析法对序列进行去噪。利用奇异谱分析法进行滤波时，不需要任何原始序列中噪声的统计特性，仅需知道原始序列大概的信噪比即可。而在消除海面高序列仪器噪声实验中，并不知道任何仪器误差的先验统计特性，且

收稿日期：2014-03-01；**修订日期**：2015-05-24
基金项目：中国科学院国家空间科学中心五个重点培育方向(Y32114FB9S)。
作者简介：叶　沛(1988－)，男，贵州贵阳人，硕士研究生，主要从事高度计定标方面的研究。E-mail：yepei1988@163.com。
通讯作者：许　可(1967－)，男，辽宁盖州人，研究员，主要从事星载雷达高度计系统技术、合成孔径雷达高度计系统技术和信号处理技术方面的研究。E-mail：xuke@mirslab.cn。

仪器的绝对误差量与真实的测量值相比较小，此时若进行最小二乘滤波，由于没有任何仪器误差的先验信息，可能因此带来额外的误差。因而可用奇异谱分析的方法对混有仪器噪声的海面高序列进行降噪处理。

2 奇异谱分析

2.1 奇异谱分析基本概念

奇异谱分析是一种基于主成分分析的序列处理方法[3]，它仅需较少先验信息即可对原始序列进行降噪处理，分析序列结构(提取原始序列中的趋势、周期成分等)，因此广泛应用于气象或地理方面时间序列的处理[4]。

由于矩阵的特征值能够反映矩阵的信息，并且可通过这些特征值及其对应的特征向量来恢复该矩阵，大的特征值对应的是信号中能量大的组分，小的特征值则可认为对应的是信号中的噪声信息。因此，当选择用较大的特征值及其对应的特征向量来恢复矩阵时，得到的就是含有较少噪声或不含噪声的矩阵，这就是奇异谱分析去噪的基本思想。奇异谱分析分为信号的分解与重建两个阶段[5-6]。

假设混有噪声的序列 $x_i(i=1,2,\cdots,N)$ 为经过了中心化的 $1\times N$ 的序列，即序列 x_i 的均值为0。为了能够利用SSA，可通过嵌入的方法，将序列 x_i 扩展为在时间上滞后排列的 $M\times(N-M+1)$ 的空间坐标矩阵(trajectory matrix) X。由式(1)可以看出，矩阵 X 是Hankel矩阵。

$$X=\begin{bmatrix} x_1 & x_2 & \cdots & x_{i+1} & \cdots & x_{N-M+1} \\ x_2 & x_3 & \cdots & x_{i+2} & \cdots & x_{N-M+2} \\ \vdots & \vdots & & \vdots & & \vdots \\ x_M & x_{M+1} & \cdots & x_{i+M} & \cdots & x_N \end{bmatrix}=[X_1,X_2,\cdots X_{i+1},\cdots X_{N-M+1}] \tag{1}$$

其中：M 为嵌入的维数。

随后，计算空间坐标矩阵 X 的 $M\times M$ 自协方差矩阵 T_x，这是一个Toeplitz矩阵，即任何平行于主对角线上的元素均相同。

$$T_x=\begin{bmatrix} c(0) & c(1) & \cdots & \cdots & \cdots & c(M-1) \\ c(1) & c(0) & c(1) & \cdots & \cdots & \cdots \\ \cdots & c(1) & \cdots & \cdots & \cdots & \cdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ \cdots & \cdots & \cdots & \cdots & \cdots & c(1) \\ c(M-1) & \cdots & \cdots & \cdots & c(1) & c(0) \end{bmatrix} \tag{2}$$

其中：$C(j)$ 是延时为 j 时 x_i 的协方差。

$$c(j)=\frac{1}{N-j}\sum_{t=1}^{N-j}x_t x_{t+j} \quad 0\leqslant j\leqslant M-1 \tag{3}$$

由式(3)可知 T_x 是非负的。对 T_x 分解特征值，得到其特征值 $\lambda_k(k=1,2,\cdots,M)$，将这些特征值按从大到小的顺序重新排列，$\lambda_1>\lambda_2>\cdots>\lambda_M\geqslant 0$，这些特征值对应的特征向量为 E^k。

定义第 k 个主成分(Principal Component，PC) a_i^k 为原始时间序列 $x_i(i=1,2,\cdots,N)$ 在 E^k 上的投影：

$$a_i^k=X_i\cdot E^k=\sum_{j=1}^{M}x_{i+j}E_j^k \quad 1\leqslant i\leqslant M-1 \tag{4}$$

求得 a_i^k 后，即可用来恢复原始序列，得到重建序列(Reconstruction Component，RC)[7-8]：

$$x_i^k=\begin{cases} \frac{1}{M}\sum_{j=1}^{M}a_{i-j}^k E_j^k & M\leqslant i\leqslant N-M+1 \\ \frac{1}{i}\sum_{j=1}^{M}a_{i-j}^k E_j^k & 1\leqslant i\leqslant M-1 \\ \frac{1}{N-i+1}\sum_{j=i-N+M}^{M}a_{i-j}^k E_j^k & N-M+2\leqslant i\leqslant N \end{cases} \tag{5}$$

原始序列元素 x_i 下的所有RC之和即可完全重建出 x_i：

$$x_i=\sum_{k=1}^{M}x_i^k, i=1,2,\cdots,N \tag{6}$$

由于特征值能够反映序列的特征成分，也就是说，利用特征值及其对应的特征向量得到的PC能够反映序列的特征，前面的一些比较大的PC代表序列中有用信号的特征成分，而后面的一些比较小

的 PC 则主要反映噪声的影响。

此外，对序列进行奇异谱分析相当于将序列按其组成成分分解成了若干份，每个成分的能量都有一个特征值与之对应，并且特征值的大小反映了各成分能量之间的相对关系，而序列的能量是一定的，所以嵌入维数 M 越大，特征值就越小，各个特征值原始序列的分解成分就越多，各成分的能量也越小。因此，嵌入维数相当于分解组成成分的“分辨率”。

利用 SSA 对原始序列进行去噪处理时，可以截取前面比较大的特征值对应的 PC 来重建原始序列，从而达到去噪的目的。设利用前 p 个 PC 恢复的序列 $\hat{x}_i$ 为：

$$\hat{x}_i = \sum_{k=1}^{p} x_i^k, i = 1,2,\cdots,N \tag{7}$$

因此，利用奇异谱分析的方法对序列进行去噪处理的关键在于选择合适的嵌入维数 M 和截断长度 p。

2.2　嵌入维数及截取长度的选取

设序列 $Y_L^{(1)}$ 和 $Y_L^{(2)}$ 分别是对原始序列进行奇异谱分析时按其组成成分分解后得到的表示其两个不同成分的序列。这两个序列的加权相关系数为 $\rho_{12}^{(w)}$。通常利用加权相关系数 $\rho_{12}^{(w)}$ 来评价两个不同成分的序列间的分离度。

$$\rho_{12}^{(w)} = \frac{(Y_L^{(1)}, Y_L^{(2)})_w}{\| Y_L^{(1)} \|_w \| Y_L^{(2)} \|_w} \tag{8}$$

其中：$\| Y_L^{(i)} \|_w = \sqrt{(Y_L^{(i)}, Y_L^{(i)})_w}$；$(Y_L^{(i)}, Y_L^{(j)})_w = \sum_{k=1}^{N} w_k y_k^{(i)} y_k^{(j)}, (i,j=1,2)$；$w_k = \min\{k, M, N-k+1\}$，$\left(k \leqslant \dfrac{N}{2}\right)$。

式(8)中 $\rho_{12}^{(w)}$ 越小，表明两个序列的加权正交性越好，它们间的分离度也越好。若 $\rho_{12}^{(w)} = 0$，则表示两个序列完全加权正交，它们是“分离”的[9]。

选取合适嵌入维数是为了使奇异谱分解时得到的各个独立的组成成分序列间的分离度尽可能高，以达到将序列中的有用信息完全从噪声中分离出来从而对序列去噪的目的。在 DGPS 浮标得到的海面高序列中，通常还包含了潮汐分量等，这在高度计定标实验中是有用的。由于嵌入维数相当于分解得到的组成成分的“分辨率”，若选择的嵌入维数过小，则潮汐分量等的组成成分有可能与仪器噪声的组成成分部分或完全混叠，使得在对海面高序列去噪时将潮汐分量等有用信息也消除了，导致定标结果的准确度和精度变差。Hossein[10] 在数学上证明了嵌入维数 M 最大只能取到 $\left[\dfrac{N+1}{2}\right]$（[·]为·的整数部分），且当 $M = \left[\dfrac{N+1}{2}\right]$ 时，各成分之间的分离度最高。

p 的选择和噪声信号与有用信号之间能量的相对大小有关。若 p 选的合适，则前 p 个主成分表示有用信号，而 p 以后的主成分反映了原始序列中的噪声。由于在 DGPS 浮标海面高序列中，仪器误差对测量结果的影响较小，若将该海面高序列类比为混有仪器噪声的信号，则仪器噪声的能量远小于有用信号的能量，因此表示噪声信号的 p 以后的特征值较小。由于仪器噪声是白噪声，其能量分布较均匀，因此表示噪声信号的 p 以后的特征值变化缓慢。故利用特征值的变化率来选择截断长度 p。

定义特征值变化率：

$$r = (\lambda_k - \lambda_{k+2})/2, k = 1,2,\cdots,M-2 \tag{9}$$

式(9)中当 $r \leqslant 0.001$，即 $\lg r \leqslant -3$ 时，可认为此时 $p = k-1$ 为截断长度。

3　仿真及结果

3.1　奇异谱分析法仿真结果

考虑一个由 DGPS 浮标得到的海面高序列 SSH(t)。

$$\mathrm{SSH}(t) = \mathrm{SSH}_0(t) + n(t) \tag{10}$$

其中：$\mathrm{SSH}_0(t)$ 为真实海面高，$n(t)$ 为仪器误差。

假定真实的海面高为 $\mathrm{SSH}_0(t) = 10 + 10^{-4} \times t$，$t$ 为时间($\mathrm{SSH}_0(t)$ 单位为 m，时间 t 单位为 s)；仪器误差为 $\sigma_n = 0.1$ m，用均值为 0、方差为 σ_n^2 的高斯白噪声模拟，需要说明的是，这里仪器误差不仅包括了浮标的仪器误差，还包括了 DGPS 数据处理中所引入的误差。序列采样间隔为 1 Hz，设定采样时间为长度为 1 h，序列长度为 $N = 3\,600$，则由浮标测得的瞬时海面高原始序列如图 1 所示。原始序列的标准差 $\sigma_{ssh} = 0.1444$ m。

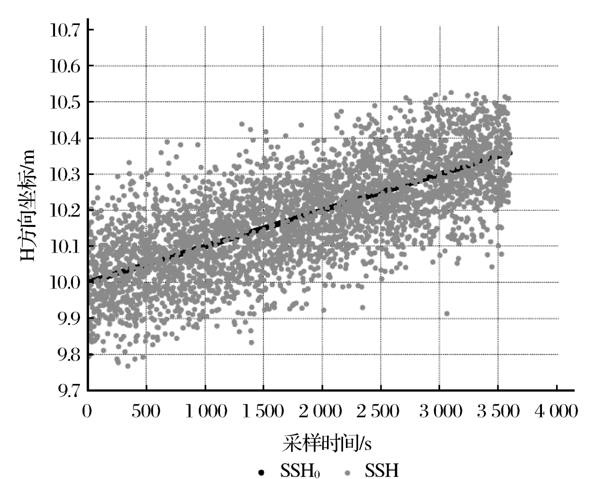


图 1　海面高序列仿真

Fig. 1　Simulated SSH series

选择嵌入维数 $M = 1\ 800$，对序列进行奇异谱分析。图2是序列分解后得到的特征值的变化率曲线。根据判定准则，$k \geqslant 9$ 时，变化率 r 即小于0.001。因此选择 $p = 8$ 作为截断长度。

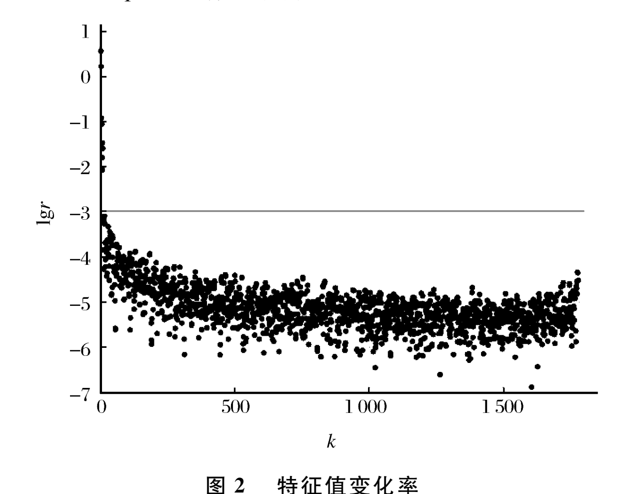


图2 特征值变化率

Fig. 2 The eigenvalues change rate

利用上述选定的 M 和 p 再对原序列进行奇异谱分析，得到去噪后的重建海面高序列如图3，它与真实海面高序列 SSH_0 基本重合。重建的序列的标准差 $\sigma_r = 0.1058$ m。

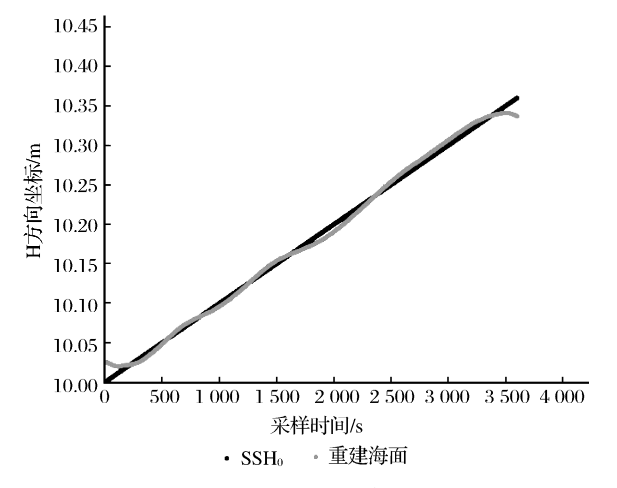


图3 重建的海面高序列

Fig. 3 The reconstructed SSH series

应用 p 以后的主成分来重建噪声信号，图4是重建的噪声序列的频谱。可以看出，其功率在整个频段上分布均匀，符合白噪声的频谱分布特性。图5是重建噪声的幅度分布直方图，其幅度符合高斯分布。重建仪器噪声的标准差 $\sigma_m = 0.09958$ m，均值 $m = -5.613 \times 10^{-5}$ m，基本符合添加的噪声的统计特性。

因此，从仿真中可以看出，用奇异谱分析的方法能够较好地将序列中的有用信号和噪声信号分离出来，其去噪效果较好。

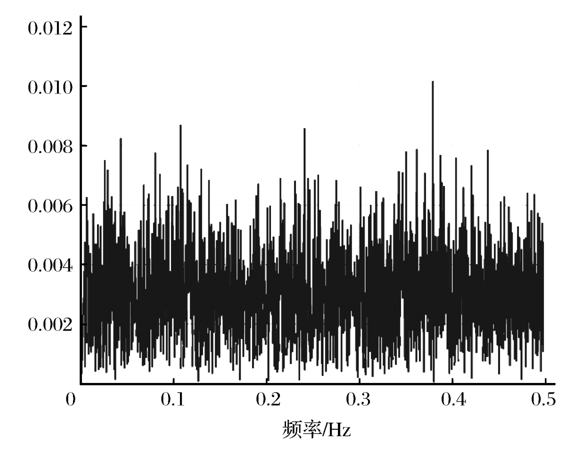


图4 重建的噪声序列频谱

Fig. 4 The frequency spectrum of reconstructed noise series

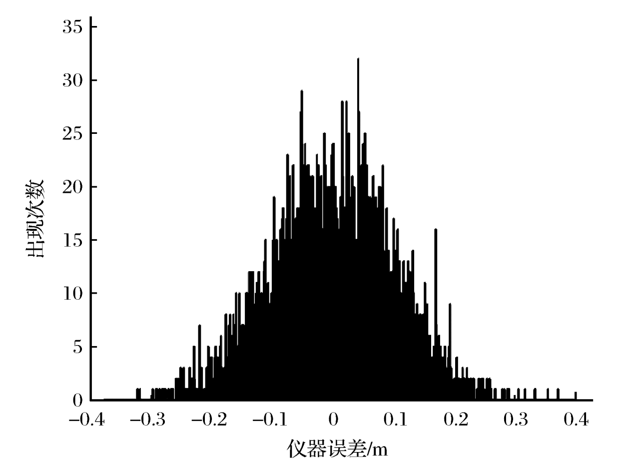


图5 仪器噪声幅度分布直方图

Fig. 5 The distribution histogram of instrumental noiseamplitude

3.2 不同嵌入维数和截断长度对仿真结果的影响

为了进一步比较不同的 M 和 p 对数据处理结果的影响，同时为了与上节仿真结果比较，分别用固定 $p = 8$ 改变 M 和固定 $M = 1\ 800$ 改变 p 这两种方法处理上述海面高序列。表1是 $p = 8$ 时不同 M 下得到的序列的标准差，表2是 $M = 1\ 800$ 时不同 p 下得到的序列的标准差。由奇异谱分析去噪的原理可知，p 必须小于 M，且 $p \geqslant 1$，故在表1中 M 从 $M = 9$ 变化到 $M = 1\ 800$；表2中 p 从 $p = 1$ 变化到 $p = 1\ 799$。

比较表1及表2可以得到，重建海面高序列的标准差是随着 M 的增大而减小，随着 p 的增大而增大的；重建噪声序列的标准差是随着 M 的增大而越接近仿真中设定的噪声标准差，随着 p 的增大而越远离设定的噪声标准差的。虽然表2中 $p = 2$ 时重建海面高序列的标准差最小，但其对应的重建噪声序列标准差较 $p = 8$ 时偏离较多，故在 $M = 1\ 800$ 下，$p = 2$ 不是最优的。因此，由以上分析可知，在上节仿真中选择 $M = 1800$，$p = 8$ 是较合适的。

表 1　$p=8$ 时不同 M 下得到的序列标准差

Table 1　The series std of Various M with $p=8$

M	9	500	1 000	1 500	1 800
重建海面高序列标准差	0.13566	0.10696	0.10657	0.10635	0.10581
重建噪声序列标准差	0.03096	0.09833	0.09919	0.09949	0.09958

表 2　$M=1\ 800$ 时不同 p 下得到的序列标准差

Table 2　The series std of Various p with $M=1\ 800$

p	1	8	500	1 000	1 500	1 799
重建海面高序列标准差	0.09490	0.10583	0.12874	0.13914	0.14444	0.14562
重建噪声序列标准差	0.10132	0.09955	0.06202	0.03727	0.01603	0.00284

4　利用 SSA 消除海面高序列仪器噪声

下面以某次水库 DGPS 浮标实验为例，对其进行奇异谱分析来消除序列中仪器误差的影响。由于实验当天水面无风或风速很小，因此，水面无风浪或者风浪很小，可认为浮标得到的海面高仅受到仪器误差的影响。查看浮标采集到的 GPS 数据并截取其中质量较好的一段，利用 GAMIT/GLOBK 软件中的动态处理模块 TRACK 解算，得到浮标所在位置的海面高数据。图 6 是浮标在大地坐标下高度（海面高）方向的运动轨迹，其标准差 $\sigma_0=2.88$ cm。

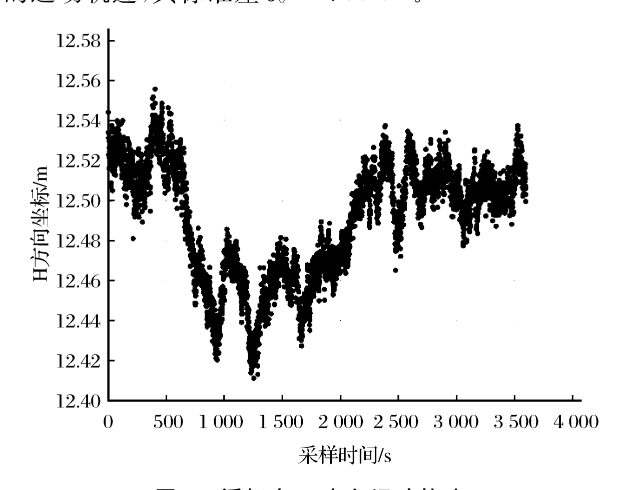


图 6　浮标在 H 方向运动轨迹

Fig. 6　The buoy H orientation movement path

选择 $M=1\ 800$ 对序列进行奇异谱分析。图 7 是分解得到的特征值的变化率。根据判定准则选择 $p=10$。

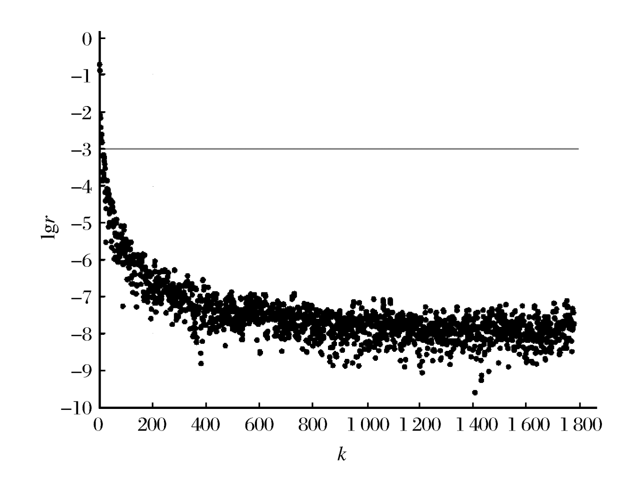


图 7　特征值变化率

Fig. 7　The eigenvalues change rate

利用上述选定的参数，再次对原序列进行奇异谱分析处理，得到的重建序列如图 8 所示，与原序列相比，重建的海面高序列较平滑。其标准差 $\sigma_e=2.69$ cm，在 RSS 意义上，标准差较处理前改善了约 1.03 cm。

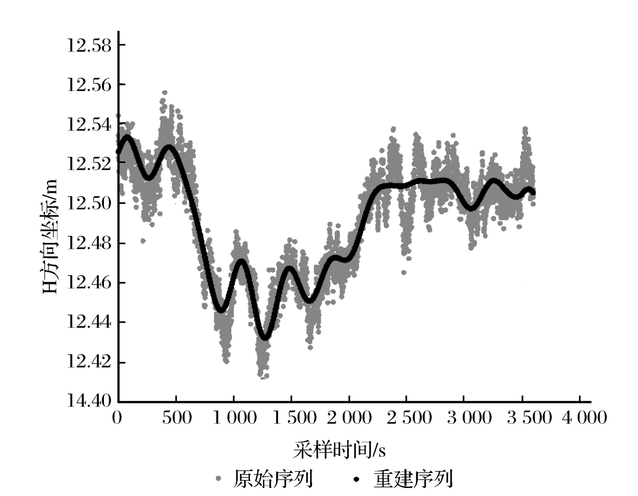


图 8　重建的海面高序列

Fig. 8　The reconstructed SSH series

5　结　语

在高度计绝对定标实验中，利用 DGPS 浮标进行现场海面高测量得到的海面高序列中，通常包括了仪器误差造成的影响。由于在浮标布放期间，仪器误差

是白噪声，因此不能通过低通滤波的方式完全消除仪器误差对海面高序列的影响。奇异谱分析是一种基于主成分分析的序列处理方法，仅需较少的先验信息即可对序列进行降噪处理。本文研究了奇异谱分析的基本理论及如何选择合适的嵌入维数和截断长度，并通过仿真验证了奇异谱分析方法及嵌入维数和截断长度选取准则的有效性。随后，利用奇异谱分析处理了某次水库 DGPS 浮标实验的海面高序列，处理结果在精度上有一定的改善。因此奇异谱分析法在消除海面高序列仪器噪声中较为适用。

参考文献(References)：

[1] Chrstopher S W. Satellite Altimeter Calibration and Validation Using GPS Buoy Technology [D]. Tasmania：Centre for Spatial Information Science，University of Tasmania，Australia，2005.

[2] Zhang Yong，Lü Daren. Vicarious Calibration of WindSat Polarimetric Channels based Sea Surface Background[J]. Remote Sensing Technology and Application，2014，29(5)：727-734.[张勇，吕达仁. 一种基于海面背景的 WindSat 极化通道替代定标方法研究[J]. 遥感技术与应用，2014，29(5)：727-734.]

[3] Schoellhamer D H. Singular Spectrum Analysis for Time Series with Missing Data[J]. Geophysical Research Letters，2001，28(16)：3187-3190.

[4] Alonso F J，Castillo J，Pintado P. Application of Singular Spectrum Analysis to the Smoothing of Raw Kinematic Signals [J]. Journal of Biomechanics，2005，38：1085-1092.

[5] Shu H C，Chin H L. Application of Singular Spectrum Analysis to Structural Monitoring and Damage Diagnosis of Bridges [J]. Structure and Infrastructure Engineering，2013：1-20.

[6] Xie J G，Ji Z P. Singular Spectrum Analysis for Tropical Cylone Landing in Guangdong[J]. Journal of Tropical Meteorology，2003，9(2)：201-206.

[7] Wu Hongbao. Singular Spectrum Analysis Combined with Maximum Entropy Method[J]. Gansu Meteorology，2000，18(1)：1-5.[吴洪宝. 奇异谱分析——最大熵预报方法[J]. 甘肃气象，2000，18(1)：1-5.]

[8] Liu Yuanfeng，Zhao Mei. De-noising of Chaotic Time Series based on Singular Spectrum Analysis[J]. Journal of Shanghai Jiaotong University，2003，37(5)：778-780.[刘元峰，赵枚. 基于奇异谱分析的混沌序列降噪[J]. 上海交通大学学报，2003，37(5)：778-780.]

[9] Hossein H. Singular Spectrum Analysis：Methodology and Comparison[J]. Journal of Data Science，2007，5：239-257.

[10] Hossein H，Rahim M，Mohammad Z. Separability and Window Length in Singular Spectrum Analysis[J]. Computers Rendus Mathematique，2011，349(17)：987-990.

Study on Denoising the Instrumental Errors of the Sea Surface Height Series Derived from Buoys

Ye Pei[1,2,3]，Xu Ke[1,2]，Xu Xiyu [1,2]

(1. *The CAS Key Laboratory of Microwave Remote Sensing*，*Center for Space Science and Applied Research*，*Chinese Academy of Sciences*，*Beijing* 100190，*China*；
2. *Center for Space Science and Applied Research*，*Chinese Academy of Sciences*，*Beijing* 100190，*China*；
3. *University of Chinese Academy of Sciences*，*Beijing* 100049，*China*)

Abstract：In absolute altimeter calibration campaigns，the Sea Surface Height (SSH) series derived from the in-situ Differential GPS (DGPS) buoys always include high frequency errors such as wind waves and instrumental noises. The former can be removed by low pass filters ，while the latter can not，because the noises are white and can spread in the whole frequency band of interest. In this paper，the Singular Spectrum Analysis (SSA) method was applied to solve the problem，and select criteria of the embedding dimension and intercepting length. Simulations based on this innovative method were executed，the effectiveness in removing the instrumental noise and the selction criteria was validated. Finally，SSH series derived from an in-situ DGPS buoy experiment were processed by this method，and the study suggested that the processed series were smoother compared to the original series，and had some improvements in accuracy. Therefore，the validity of the SSA method was demonstrated.

Key words：Altimetric calibration；SSH series；Instrumental errors；Singular Spectrum Analysis(SSA)；De-noise